Transfer Learning Approach for Occupancy Prediction in Smart Buildings

Mohamad Khalil
School of Engineering
Newcastle University
Newcastle, Uk
m.khalil2@newcastle.ac.uk

Stephen McGough
School of Computing
Newcastle University
Newcastle, UK
stephen.mcgough@ncl.ac.uk

Zoya Pourmirza
School of Engineering
Newcastle University
Newcastle, UK
zoya.pourmirza@ncl.ac.uk

Mehdi Pazhoohesh
Faculty of Technology
De Montfort University
Leicester, UK
mehdi.pazhoohesh@dmu.ac.uk

Sara Walker
School of Engineering
Newcastle University
Newcastle, UK
sara.walker@ncl.ac.uk

2021 12th International Renewable Engineering Conference (IREC) | 978-1-6654-2548-3/20/$31.00 ©2021 IEEE | DOI: 10.1109/IREC51415.2021.9427869

***Abstract*— Accurate occupancy prediction in smart buildings is a key element to reduce building energy consumption and control HVAC systems (Heating – Ventilation and– Air Conditioning) efficiently, resulting in an increment of human comfort. This work focuses on the problem of occupancy prediction modelling (occupied / unoccupied) in smart buildings using environmental sensor data. A novel transfer learning approach was used to enhance occupancy prediction accuracy when the amounts of historical training data are limited. The proposed approach and models are applied to a case study of three office rooms in an educational building. The data sets used in this work are actual data collected from the Urban Sciences Building (USB) in Newcastle University. The results of the proposed transfer learning approach have been compared with the models from Support Vector Machine and Random Forest algorithms. The final results demonstrate that the most accurate model in this study to predict occupancy status was produced by stacked Long-Short-Term-Memory with a transfer learning framework.**

Keywords—(Occupancy Prediction, Machine Learning, Transfer Learning, Deep Learning, Smart Buildings)

Nomenclature

ML	Machine learning
DL	Deep learning
SVM	Support vector machine
RF	Random forest
LSTM	Long-Short-Term-Memory
DT	Decision trees
TL	Transfer learning
HVAC	Heating, ventilation and air conditioning
MLP	Multi-Layer Perceptron
BPN	Back-Propagation Network
LDA	Linear discriminant analysis
CNN	Convolutional neural network
KNN	K-Nearest Neighbour
AMI	Advanced metering infrastructure
CART	Classification and Regression Trees
PIR	Passive infrared sensor
ReLu	Rectified linear unit
D	Domain
X	Feature space
$P(X)$	Marginal probability distribution
T	Task
Y	Label space
$F(\cdot)$	Objective predictive function

I. INTRODUCTION

In buildings the term of occupancy information refers to occupants' presence or absence and their movements.

Accurate real-time occupancy information in smart buildings can help HVAC (Heating – Ventilation and– Air Conditioning) control systems to optimize their usage and minimize building energy consumption [1]. In the past years, occupancy prediction modelling in buildings has been studied by many researchers using different types of data such as air temperature, sound, door status, relative humidity, camera, motion and light.

Recently, Machine Learning (ML) and Deep Learning (DL) models are becoming more popular in the applications of building energy such as forecasting building energy consumption and occupancy prediction [2,3].

The present work focuses on the challenging task of occupancy prediction modelling (occupied / unoccupied) when the amounts of historical training data are limited by using a transfer learning framework.

The rest of this paper is organized as follows: Section II, introduces a brief summary on ML and DL applications in related works. Section III, discusses data collection and data pre-processing steps in this paper. Section IV, introduces the proposed approach and algorithms. Section V, shows the results. Finally, Section VI, presents the conclusion.

II. RELATED WORK

Occupancy prediction modelling (occupied / unoccupied) can be categorized as a classification task, which means that each row in a data set assigns to a class.

Recently, researchers have used different approaches of ML and DL to predict occupancy status (absence / presence). The reviewed literature has been categorized and presented in Table I based on the following: year, classification algorithms, buildings type, type of features used, size of historical data and reported accuracy.

Occupancy prediction (occupied / unoccupied) in a domestic environment was presented by Vafeiadis et al., (2017) in [4]. The authors used decision trees (DT), support vector machine (SVM), back-propagation network (BPN), and random forest (RF), along with the AdaBoost algorithm to predict occupancy status. These ML models used data from water and energy consumption. It was found that RF and DT outperformed other classification techniques with respect to the accuracy.

978-1-6654-2548-3/21/$31.00 ©2021 IEEE

Three statistical classification algorithms which are linear discriminant analysis (LDA), classification and regression trees (CART) and RF were used by Candanedo and Feldheim (2016) in [1] to predict occupancy in an office environment. Their results have shown that appropriate selection of features can significantly improve the prediction performance.

Sensor data from light, carbon dioxide (CO_2), passive infrared motion (PIR), sound and electrical current were used to train a decision tree model by Hailemariam , Goldstein, Attar, and Khan (2011) [5].

Wang, Chen, and Hong (2018) in [6] used the following ML models: k-nearest neighbor (KNN), SVM and BPN. These models were trained with environmental and Wi-Fi data to classify occupancy pattern. It was found that fused data sets can improve the reliability of these techniques.

A deep neural network to predict occupancy status was presented by Abedia and Jazizadehb (2019) in [7], and the data used in this study were collected from two cost-effective sensors: doppler radar sensors (DRS) and infrared thermal array (ITA).

Feng, Mehmani, and Zhang (2020) in [8] used advanced metering infrastructure (AMI) data to train a DL model which consists of convolutional neural network (CNN) and bidirectional LSTM. The average reported accuracy of this model was around 90%.

A study of real-time occupancy detection in an office environment using CO_2 data was presented by Zhou. et al., (2019) in [9]. The ML model used in this study was gcForest.

However, according to these studies ML and particularly DL models rely on large amounts of historical training data to provide satisfactory prediction and forecasting accuracy, but sometimes this data are difficult to collect or very expensive. In fact, a group of newly monitored buildings do not have enough amounts of historical observations and measurements to predict occupancy status accurately. In light of this challenge, a novel transfer learning approach for occupancy prediction modelling (occupied / unoccupied) in an office environment is presented to obtain a desirable accuracy when the amounts of historical training data are limited.

Transfer learning is a subset of ML. The main idea of transfer learning is that the knowledge that has already been learned from a related source domain can be transferred to improve model performance in a target domain [10]. Gao, Ruan, Fang, and Yin (2020) in [11] utilized two DL models which are a sequence-to-sequence and a two-dimensional convolutional neural network with transfer learning to predict building energy consumption. Their results indicate that transfer learning could improve the prediction accuracy.

The objective of this study is to determine whether transfer learning approach could improve the occupancy prediction performance of target rooms with limited data by transferring the knowledge learned from a similar room with enough historical data.

III. CASE STUDY

The Urban Science Building (USB) studied in this paper is a university educational building situated in City-Of-Newcastle in the United Kingdom, Fig. 1 provides a picture showing the building. USB is one of the largest repositories of publicly held data in the UK. This building has six floors and it mainly comprises of office and teaching rooms. The office hours of this building approximately run from 08:00 to 18:00 every weekday. Three office rooms (A, B and C) within the third and fourth floors were selected in our experiments. These rooms run a similar business schedule throughout the year. The rooms differ in the size of the area covered (room A: 31 m^2, room B: 30 m^2, room C: 15m^2).

A. Data collection

This paper used data collected from environmental sensors such as temperature, relative humidity and CO_2 , in addition to motion sensor. This use of environmental sensor data can avoid privacy concerns of using internal cameras for real-time occupancy detection.

Table I. Comparison of the reviewed studies

Reference No	Year	Classification algorithms used	Environment type	Data type	Size of historical data	Accuracy
[1]	2016	LDA, CART, RF	Office	Environmental: temp, humidity, light, CO_2, humidity ratio	1 month	Ranging from 32.68% to 99.31%
[4]	2017	SVM, RF, DT, BPN, AdaBoost	Domestic	Smart meter: Lights, refrigerator, TV, Washing Machine, Dryer, Cold Water – Kitchen, Hot Water Kitchen, Dishwasher - Water, Washing Machine - Water	1 month	Ranging from 75.87% to 80.53%
[5]	2011	DT	Office	Environmental: light, CO_2, PIR, electrical current, sound	Seven day	Ranging from 81.01% to 98.44%
[6]	2018	BPN, KNN, SVM	Office	Environmental (indoor air temperature, RH, and CO_2), Wi-Fi and Camera.	Nine days	NA
[7]	2019	DNN	Office	Data from DRS and ITA sensors	NA	Ranging from 98.9% to 99.9 %
[8]	2020	CNN-BiLSTM, KNN, SVM, RF, multi-layer perceptron (MLP), AdaBoost.	Domestic	Electricity consumption and occupancy	NA	Ranging from 52.70% to 98.4%
[9]	2019	gcForest	Office	Environmental: CO_2	1 month	83.3 %

Data collection period can be found in Table II. Fig. 2 shows environmental measurements of room (C) between 01st Feb 2020 and 01st Mar 2020.

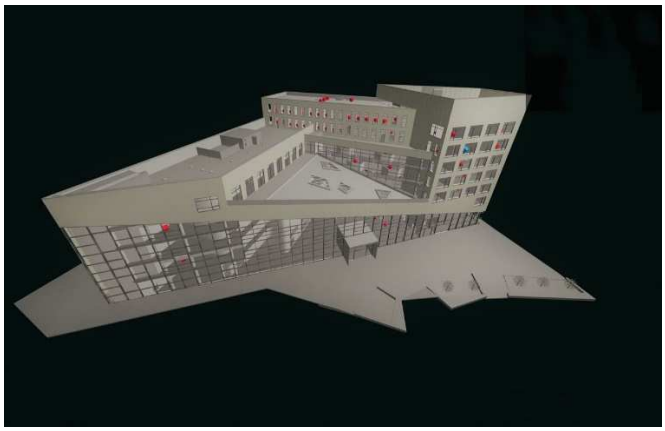

Fig. 1. USB building

Table II. Data collection

Room ID	Period
Room (A)	Data collected between January 2018 and March 2020
Room (B)	Date collected between January 2020 and March 2020
Room (C)	Date collected between January 2020 and March 2020

B. Data preprocessing

Data preprocessing is a fundamental step in ML and DL models to improve the model's accuracy and performance [12]. In the scope of this study, five data preprocessing steps were used. These steps are: data re-sampling, feature engineering, min-max, z-score and duplicate records elimination.

- In ML, resampling refers to the process of converting time series data from one frequency to another. All the features of our data sets were resampled to regular frequency of 5 mins interval to maintain stability and consistency.
- In terms of features engineering, seven time-based features were extracted from each time stamp. These features are year, month, day, hour, minute, week-day (binary feature denoting whether it is weekday or weekend) and work pattern (denoting whether it is business hour or non-business hour).
- Min-Max normalization is a very important method in ML and DL [11]. This normalization technique was applied in this study to rescale the features.
- To detect and eliminate outliers, z-score technique was calculated for each numerical feature in our data sets. Any z-score value greater than 3 or less than 3 considered to be an outlier and thus to be removed it.
- Duplicate records have been removed.

IV. Methodology

In this section, the proposed approach is demonstrated by using two DL models which are a stacked long-short-term-memory (LSTM) and a sequential deep model, and then apply transfer learning on top of them. In addition, two supervised ML techniques which are SVM and RF have been selected to assess the effectiveness of our proposed approach. These supervised models are well suited for classification problems of small-sized data sets, and they work well with non-linearly separable data.

The models were implemented using the Python 3.8.3 programming language, keras deep learning and scikit-learn library.

A. Long-short-term-memory (LSTM)

LSTM is a type of recurrent neural network. This can usually be used to process sequential data and solve time series prediction problems [13]. This type of neural networks can handle time dependency through feeding back the output of a layer at time (t) to the input of the same layer at time (t + 1). In the scope of this study, stacked LSTM network was utilized.

The architecture of this model and the optimal hyperparameters set for building stacked LSTM are listed in table III and table IV respectively.

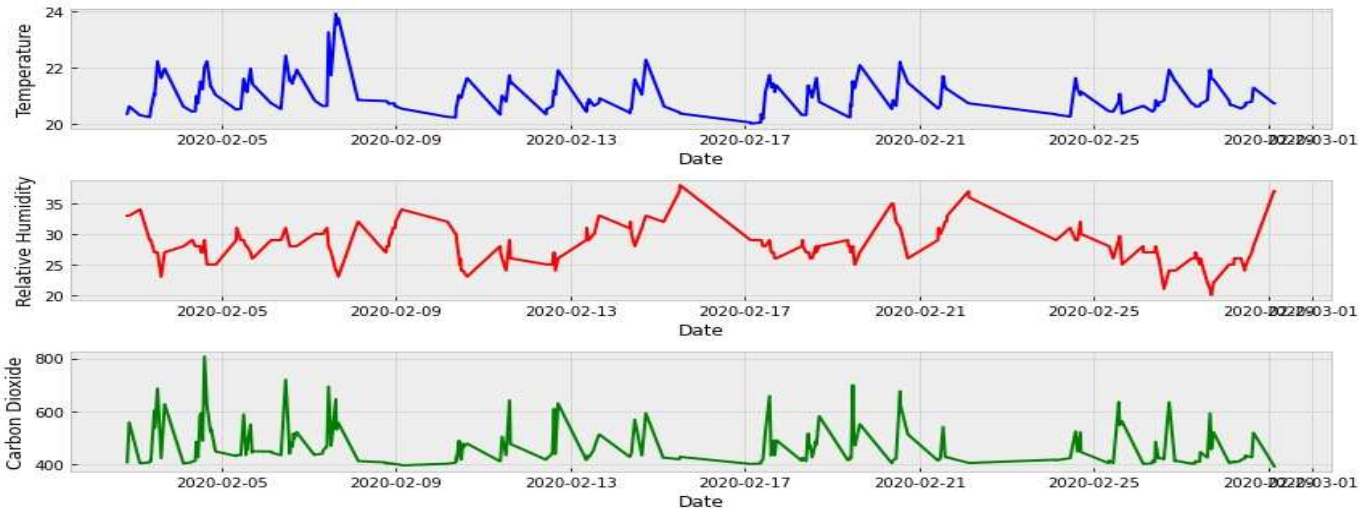

Fig. 2. Room (C)' environmental measurements in February 2020

B. Sequential deep model

Sequential model is the simplest way to build a deep neural network. This model usually consists of plain stack of layers. In this study, three hidden layers were used to build this model. The architecture of this model and the optimal hyperparameters set can be found in table V and table VI respectively.

C. Random Forest (RF)

This algorithm can be used for classification and regression tasks. RF is an ensemble technique constructed from multiple decision trees, where each tree of them is using a different bootstrap sample of training data [14]. Add to that RF is one of the ways in ML to avoid overfitting in traditional tree-based models [15]. Walker, Khan, Katic, Maassen, and Zeiler (2020) in [16] applied RF, SVM, boosted-tree and artificial neural network (ANN) to predict electricity demand in building. Their results have shown that RF and boosted-tree outperformed SVM and ANN with respect to the performance.

For optimizing RF model, this study takes the hyperparameters found in table VII into consideration.

D. Support vector machine (SVM)

SVM can be used to solve linear and non-linear classification and regression problems in addition to its ability to perform well on limited amounts of training data [17]. The mechanism of this algorithm in a classification task is to find an optimal hyperplane that maximizes the margin between classes of a data set [18]. For building a SVM model, the polynomial function is chosen as a kernel function in this study after preliminary analysis that found the data set is non-linearly separable.

E. Transfer learning approach

Nowadays, ML and DL prediction models have been widely used in variety of domains such as natural language processing and computer vision. However, the amounts of historical training data needed to provide satisfactory prediction results are beyond what most researchers and industries can get. To address this limitation, transfer learning has emerged as a promising concept in ML [19]. Domain and task are two fundamental concepts in transfer learning.

A domain D is composed of a feature space X and a marginal probability distribution $P(X)$, $D = [X, P(X)]$ [20]. In the field of this study, the feature space includes CO_2, relative humidity, temperature, and other features. A task T can be defined as $T = [Y, F(\cdot)]$, where Y is a label space and $F(\cdot)$ an objective predictive function [20]. In this study, Y is the target occupancy feature (occupied / unoccupied) to be predicted. Fig. 3 shows the structure of the transfer learning approach used in this study.

The proposed approach is simple and includes two steps. The first step is to train the DL models that are stacked LSTM and sequential deep model on a large source data set (i.e., room A) that contains two years of historical training data. The second step is to fine-tune the same models on the target data sets (i.e., room B and room C) which contain two months of historical training data. The experiments setup can be found in Table VIII.

Table III. stacked LSTM model architecture

Layer	Description
LSTM	Neurons = 30, input shape = (10,1), return sequence = True
LSTM	Neurons = 30
Dense output	Neurons = 1

Table IV. Hyperparameters for stacked LSTM

Hyperparameter	Setting
Activation function	ReLu
Kernel initializer	Truncated normal
Optimizer	Adam
Loss	Binary cross entropy
Output Activation function	Sigmoid

Table V. Sequential model architecture

Layer	Description
Dense	Neurons = 30
Dense	Neurons = 30
Dense	Neurons = 30
Dense output	Neurons = 1

Table VI. Hyperparameters for sequential model

Hyperparameter	Setting
Activation function	ReLu
Optimizer	Adam
Batch size	32
Loss	Binary cross entropy
Output Activation function	Sigmoid

Table VII. Hyperparameters for RF

Hyperparameter	Setting
Criterion	Gini
Max depth	4
Max feature	Auto
N-estimator	500

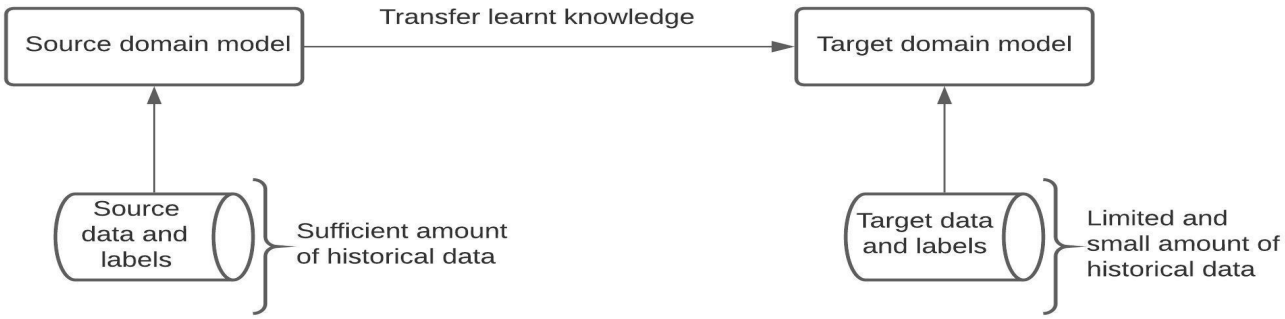

Fig. 3. Transfer learning approach

Table VIII. Experiments setup

Experiment ID	Source domain/room	Target domain/room
Experiment (1)	Room A	Room B
Experiment (2)	Room A	Room C

The following describe the details of the models used in this study:

- Model 1: RF trained with two months of data from the target room.
- Model 2: SVM trained with two months of data from the target room.
- Model 3: sequential deep model without transfer learning trained with two months of data from the target room.
- Model 4: stacked LSTM without transfer learning trained with two months of data from the target room.
- Model 5: sequential deep model with transfer learning trained with two years of data from the source room (A) and two months of data from the target room used as a fine-tuning data set.
- Model 6: stacked LSTM with transfer learning trained with two years of data from the source room (A) and two months of data from the target room used as a fine-tuning data set

V. Results and Evaluation

A. Evaluation criteria

In this study, the classification accuracy has been selected and used as a measure to evaluate the prediction performance. Classification accuracy refers to the total number of records that are correctly classified by a classification algorithm. The accuracy results of room (B) and room (C) are plotted in Fig. 4 and Fig. 5 respectively, and full details of the accuracy results are listed in Table IX.

B. Comparison between transfer learning approach and supervised learning techniques

As can be seen for the Fig. 4. and Fig. 5. the stacked LSTM with transfer learning emerges as the winner in each experiment in terms of the accuracy. Sequential deep model with transfer leaning also has the second best accuracy. Supervised ML techniques SVM and RF performed worse than all the transfer learning approaches with respect to the accuracy. Overall , these results indicate the necessity to use and apply a transfer learning approach to improve the occupancy prediction performance.

C. Comparison between transfer learning approach and DL models without transfer learning

DL models without a transfer learning in our experiments have achieved accuracy between 62% and 66% as can be shown in Table IX. In summary, these results show that applying transfer learning on top of the DL models can improve the prediction accuracy in this work.

Table IX. Results

Model Number	Room B	Room C
Model 1	64 %	62 %
Model 2	66 %	64 %
Model 3	65 %	62 %
Model 4	66 %	63 %
Model 5	67 %	66 %
Model 6	71 %	69 %

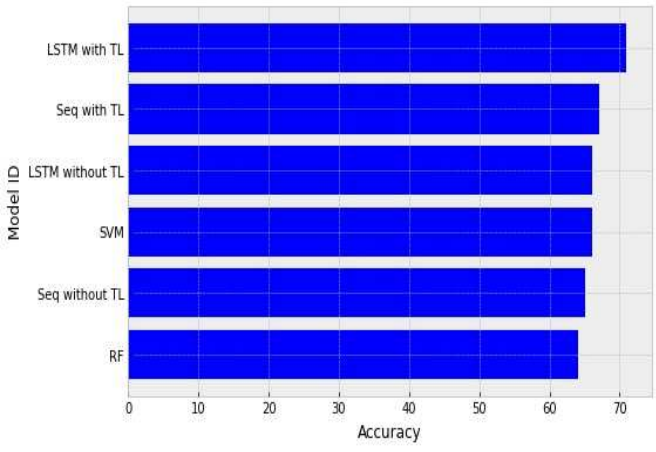

Fig. 4. A comparison of the models in room (B)

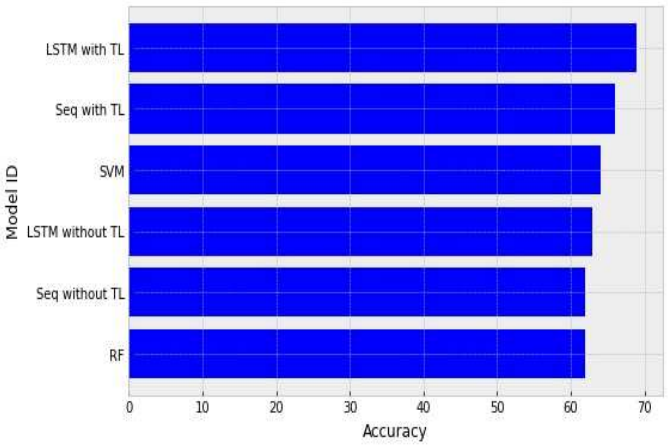

Fig. 5. A comparison of the models in room (C)

VI. CONCLUSION

In this study, a transfer learning approach was used to predict occupancy status in an educational building via the use of environmental sensor data. The proposed approach used two deep learning models which are stacked LSTM and sequential deep model, and then applied transfer learning on top of them. The results show that a transfer learning approach is a reliable and suitable framework for the challenging task of occupancy prediction modelling.

Additionally, two supervised machine learning models were utilized which are SVM and RF to evaluate the need to use additional data from similar rooms. The research results indicate that, it is useful to apply a transfer learning as this approach can improve the occupancy prediction performance when the amounts of historical training data are limited.

As a future research, the plan is to focus on applying different type and structure of deep learning models. In addition, further investigations are required to use and apply a transfer learning approach for other types of buildings such as the domestic.

ACKNOWLEDGMENT

This work has been funded and supported by Newcastle University and Engineering and Physics Science Research Council (EPSRC) programme grant EP/S016627/1. In this work we used the API provided by the urban observatory link to the website: https://api.usb.urbanobservatory.ac.uk/

REFERENCES

[1] L. M. Candanedo and V. Feldheim, "Accurate occupancy detection of an office room from light, temperature, humidity and CO 2 measurements using statistical learning models," Energy and Buildings, vol. 112, pp. 28-39, 2016, doi: 10.1016/j.enbuild.2015.11.071.

[2] M. Bourdeau, X. q. Zhai, E. Nefzaoui, X. Guo, and P. Chatellier, "Modeling and forecasting building energy consumption: A review of data-driven techniques," Sustainable Cities and Society, vol. 48, 2019, doi: 10.1016/j.scs.2019.101533.

[3] C. Robinson et al., "Machine learning approaches for estimating commercial building energy consumption," Applied Energy, vol. 208, pp. 889-904, 2017, doi: 10.1016/j.apenergy.2017.09.060.

[4] T. Vafeiadis et al., "Machine Learning Based Occupancy Detection Via The Use of Smart Meters," presented at the 2017 International Symposium on Computer Science and Intelligent Controls (ISCSIC), 2017.

[5] E. Hailemariam, R. Goldstein, R. Attar, and A. Khan, "Real-time occupancy detection using decision trees with multiple sensor types," presented at the Proceedings of the 2011 Symposium on Simulation for Architecture and Urban Design, Boston, Massachusetts, 2011

[6] W. Wang, J. Chen, and T. Hong, "Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings," Automation in Construction, vol. 94, pp. 233-243, 2018, doi: 10.1016/j.autcon.2018.07.007.

[7] M. Abedia and F. Jazizadehb, "Deep-learning for Occupancy Detection Using Doppler Radar and Infrared Thermal Array Sensors," presented at the International Symposium on Automation and Robotics in Construction (ISARC 2019), 2019.

[8] C. Feng, A. Mehmani, and J. Zhang, "Deep Learning-Based Real-Time Building Occupancy Detection Using AMI Data," IEEE Transactions on Smart Grid, vol. 11, no. 5, pp. 4490-4501, 2020, doi: 10.1109/tsg.2020.2982351.

[9] Y. Zhou, Z. J. Yu, J. Li, Y. Huang, and G. Zhang1, "A data mining model for building occupancy estimation based on deep learning methods," presented at the IOP Conference Series: Materials Science and Engineering, 2019.

[10] O. Day and T. M. Khoshgoftaar, "A survey on heterogeneous transfer learning," Journal of Big Data, vol. 4, no. 1, 2017, doi: 10.1186/s40537-017-0089-0.

[11] Y. Gao, Y. Ruan, C. Fang, and S. Yin, "Deep learning and transfer learning models of energy consumption forecasting for a building with poor information data," Energy and Buildings, vol. 223, 2020, doi: 10.1016/j.enbuild.2020.110156.

[12] M. Pazhoohesh, Z. Pourmirza, and S. Walker, "A Comparison of Methods for Missing Data Treatment in Building Sensor Data," presented at the IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE), 2019

[13] I. Sulo, S. R. Keskin, G. Dogan, and T. Brown, "Energy Efficient Smart Buildings: LSTM Neural Networks for Time Series Prediction," presented at the 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019.

[14] C. E. Kontokosta and C. Tull, "A data-driven predictive model of city-scale energy use in buildings," Applied Energy, vol. 197, pp. 303-317, 2017, doi: 10.1016/j.apenergy.2017.04.005.

[15] A. M. Prasad, L. R. Iverson, and A. Liaw, "Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction," Ecosystems, vol. 9, no. 2, pp. 181-199, 2006, doi: 10.1007/s10021-005-0054-1.

[16] S. Walker, W. Khan, K. Katic, W. Maassen, and W. Zeiler, "Accuracy of different machine learning algorithms and added-value of predicting aggregated-level energy performance of commercial buildings," Energy and Buildings, vol. 209, 2020, doi: 10.1016/j.enbuild.2019.109705.

[17] S. Seyedzadeh, F. P. Rahimian, I. Glesk, and M. Roper, "Machine learning for estimation of building energy consumption and performance: a review," Visualization in Engineering, vol. 6, no. 1, 2018, doi: 10.1186/s40327-018-0064-7.

[18] Q. Chai, H. Wang, Y. Zhai, and L. Yang, "Using machine learning algorithms to predict occupants' thermal comfort in naturally ventilated residential buildings," Energy and Buildings, vol. 217, 2020, doi: 10.1016/j.enbuild.2020.109937.

[19] T. M. Lai, T. Bui, N. Lipka, and S. Li, "Supervised Transfer Learning for Product Information Question Answering," presented at the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 2018.

[20] C. Fan et al., "Statistical investigations of transfer learning-based methodology for short-term building energy predictions," Applied Energy, vol. 262, 2020, doi: 10.1016/j.apenergy.2020.114499.